


Name:			
Enrolment No:			
UPES End Semester Examination, May 2024			
Program: B.Tech CSE- BD Course: BigData Analysis Course Code: CSBD 3009		Semester: VI Time : 03 hrs. Max. Marks: 100	
Instructions: Explain in short. (60-70 words)			
SECTION A (5Qx4M=20Marks)			
S. No.		Marks	CO
Q 1	Discuss the four computing resource of BigData.	4	CO1
Q 2	Differentiate in between nominal and ordinal types of data.	4	CO 2
Q 3	Deliberate statistical significance in data analytics.	4	CO 3
Q 4	Explain the concept of hive bucket and its functioning.	4	CO 3
Q 5	Differentiate in between Data Munging and ETL.	4	CO 4
SECTION B (4Qx10M= 40 Marks)			
Instruction: Write brief notes. (100-150 words)			
Q 6	Explain the conception of structured, semi structured data with suitable examples.	10	CO1
Q7	Discuss the functioning of worker and driver node in spark cluster.	10	CO1
Q8	Converse regarding CTE and CTAS in terms of table creation in HIVE. OR Discuss the role of Driver and Thrift server component in HIVE architecture.	10	CO2
Q9	Discourse the concept of query data and managing output in HIVE.	10	CO2
SECTION-C (2Qx20M=40 Marks)			
Instruction: Write long answer. (Up to 350 words while explaining) Attempt any part of question no. 10 as there is an option “a” OR “b”. There is no choice for question no.11.			

Q 10	Deliberate about the use case of BigData in healthcare and finance domain. OR Discourse the significance of HIVE bucket and how it can be viewed after loading table.	20	CO4
Q 11	Converse about Rapid Miner and Rattle GUI tools used for BigData Analysis with suitable examples.	20	CO3